

# Discrete variational Hamiltonian mechanics

**S Lall and M West**

Aeronautics and Astronautics, Stanford University, Stanford, CA 94305, USA

E-mail: [lall@stanford.edu](mailto:lall@stanford.edu) and [westm@stanford.edu](mailto:westm@stanford.edu)

Received 13 October 2005

Published 24 April 2006

Online at [stacks.iop.org/JPhysA/39/5509](http://stacks.iop.org/JPhysA/39/5509)

## Abstract

The main contribution of this paper is to present a canonical choice of a Hamiltonian theory corresponding to the theory of discrete Lagrangian mechanics. We make use of Lagrange duality and follow a path parallel to that used for construction of the Pontryagin principle in optimal control theory. We use duality results regarding sensitivity and separability to show the relationship between generating functions and symplectic integrators. We also discuss connections to optimal control theory and numerical algorithms.

PACS numbers: 02.30.Hq, 02.40.–k, 02.60.Jh

Mathematics Subject Classification: 34A26, 49N15

## 1. Introduction

The theory of discrete Lagrangian mechanics and that of the corresponding variational integrators is reasonably well developed (see [18] and references therein). There has, however, been a lack of a corresponding Hamiltonian side in this discrete setting. In this paper we present a particular canonical choice for the Hamiltonian analogue of discrete Lagrangian mechanics, based on Lagrange duality and using a method analogous to that in optimal control theory. In addition, from the optimization and duality theory it is simple to understand reciprocity, symplecticity and momentum map preservation.

Having a Hamiltonian side for discrete mechanics is of interest for theoretical reasons, such as the elucidation of the relationship between symplectic integrators, discrete-time optimal control, and distributed network optimization, as well as practical reasons, such as the fact that variational integrators for degenerate Lagrangian systems may be more easily derived on the Hamiltonian side.

The Euler–Lagrange and Hamilton equations are equivalent when the Legendre transform is a global diffeomorphism (that is, the systems are hyper-regular), an assumption that gives some of the same power as the assumption of strong convexity in optimization. The relationship between Lagrangian and Hamiltonian dynamics is of particular importance when the system is not hyper-regular (that is, when it is degenerate) and so the two theories are not entirely

equivalent. Examples of this occur in point vortex models of fluid flow and quasi-geostrophic reduced models of atmospheric dynamics, and when deriving variational integrators for such systems it is important to make the appropriate choice between Lagrangian and Hamiltonian formulations [17].

There have been a number of other formulations of discrete Hamiltonian mechanics. Direct approaches on the Hamiltonian side, where one discretizes the Hamiltonian and the symplectic structure, are developed in [10] and developed and generalized much further in [19]. In this paper we use an approach based on duality, and similar approaches are standard in optimal control theory (see [1] for a standard treatment). Some of the applications of this theory to generator functions and variational integrators are explored in [5].

When dealing with optimization and variational problems such as those arising in discrete mechanics, there are two basic approaches that we can take. Either we assume convexity of our functions, in which case we can globally minimize or maximize, or we assume smoothness, in which case we can find stationary points locally. Here we choose to follow the second path, as is more traditional and generally useful in mechanics. We will, however, write ‘min’ and ‘max’, with the understanding that in the context of mechanics they should be interpreted as merely requiring local stationarity<sup>1</sup>.

In constructing the dual of optimization or variational problems we naturally obtain dual variables. Depending on the problem these have many interpretations, including Lagrange multipliers, canonical momenta, forces, pressures or prices. In this paper we will refer to them either as momentum variables or Lagrange multipliers.

## 2. Discrete Hamiltonian mechanics

In this paper, we construct a Hamiltonian approach to discrete mechanics, motivated significantly by the theory of convex duality.

Take a *discrete Lagrangian* function  $L : Q \times Q \rightarrow \mathbb{R}$ . To approximate a given continuous system, approximate the action

$$L(q(0), q(\Delta t)) \approx \int_0^{\Delta t} L(q(t), \dot{q}(t)) dt \quad (1)$$

on a solution trajectory  $q(t)$ . The *action* is

$$S(q(\cdot)) = \sum_{t=0}^{n-1} L(q(t), q(t+1)) \quad (2)$$

and rendering the action stationary gives the *Euler–Lagrange equations*

$$D_2 L(q(t-1), q(t)) + D_1 L(q(t), q(t+1)) = 0.$$

Here we use the notation  $D_i f$  to refer to the derivative of  $f$  with respect to its  $i$ th argument.

*Duality.* The above construction of discrete mechanics is well known (see [11] and [18] for historical references) and has been further developed in the context of numerical integration algorithms which have important geometric properties, including symplecticity and momentum-map preservation in the presence of symmetries.

Note that discrete Euler–Lagrange equations are simply the optimality conditions for an unconstrained optimization problem. The major stepping-stone to construct a corresponding Hamiltonian theory is to form the appropriate dual problem. This is not quite well defined,

<sup>1</sup> It is important to remember that Hamilton’s principle is really the principle of stationary action, and not of minimum action.

however, as there is no canonical dual problem associated with an unconstrained optimization problem. Indeed, even for constrained optimization problems, the dual problem one obtains depends on the particular equations used to describe the constraint manifold. This concept is well known in constrained optimization theory, and one can achieve significant efficiencies in numerical algorithms by selecting the appropriate formulation of the problem, and hence its dual. It is also possible to formulate dual problems by introducing artificial constraints, which is necessary for unconstrained problems. As we shall see, this is key to constructing a Hamiltonian version of discrete mechanics, and the choice of such constraints is very important when one considers multi-symplectic extensions and algorithms. We first provide a very brief review of Lagrange duality theory to guide us in our development. More details on convex duality theory and its implications may be found in [2].

*Lagrange duality.* Suppose the functional  $f$  is convex and the function  $h$  is affine. We have a constrained optimization problem given by

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && h(x) = 0. \end{aligned} \tag{3}$$

To construct the dual, one forms the *Lagrangian* (as the term is used in optimization)

$$K(x, p) = f(x) + p^T h(x), \tag{4}$$

and the *Lagrange dual function*  $g$

$$g(p) = \inf_x K(x, p).$$

Then, subject to additional technical conditions called *constraint qualifications*, the optimal value of

$$\sup_p g(p) \tag{5}$$

is equal to the minimum of (3); this optimization is called the *dual problem*. Note that the term *Lagrangian* in convex optimization is used to refer to the cost function augmented by the addition Lagrange multiplier (4), not the usual Lagrangian that arises in mechanics.

*Hamilton’s principle on the Hamiltonian side.* To form a corresponding Hamiltonian theory, we introduce additional variables and constraints to obtain the following optimization problem:

$$\begin{aligned} &\text{minimize} && S(q, r) = \sum_{t=0}^{n-1} L(q(t), r(t+1)) \\ &\text{subject to} && r(t) = q(t) \quad \text{for } t = 1, \dots, n-1 \\ &&& q(0) = r(0) \\ &&& r(n) = q(n). \end{aligned}$$

It is clear that this problem is identical to the original variational problem of Hamilton’s problem and both problems have the same solution, and the variables  $r(1), \dots, r(n)$  are redundant. Having augmented the system with these redundant variables, we now construct the dual problem.

Taking Lagrange multipliers  $p(t)$ , to construct the dual function we write the Lagrangian (as the term is used in optimization)

$$K(q, r, p) = \sum_{t=0}^{n-1} L(q(t), r(t+1)) + \sum_{t=1}^n p(t)^T (q(t) - r(t)) \tag{6}$$

with respect to both  $r(t)$  and  $q(t)$ . In this setting the Lagrange multipliers are known as the *canonical momenta*. In doing so, we can perform the minimization with respect to  $r(t)$  first. To do this, define the associated *right Hamiltonian*<sup>2</sup>

$$H^+(q(t), p(t+1)) = -\min_{r(t+1)} (L(q(t), r(t+1)) - p(t+1)^T r(t+1)).$$

Requiring stationarity with respect to variations in  $r(t)$  gives the *right Legendre transform*

$$p(t+1) = D_2 L(q(t), r(t+1)). \quad (7)$$

It is therefore a consequence of strong duality that along the trajectory that extremizes the action the Lagrange multiplier  $p$  satisfies the Legendre transform. The Lagrange dual function is

$$\begin{aligned} g(p) &= \min_{q,r} K(q, r, p) \\ &= \min_q S(q, p) \end{aligned} \quad (8)$$

where

$$S(q, p) = \sum_{t=0}^{n-1} (p(t+1)^T q(t+1) - H^+(q(t), p(t+1))). \quad (9)$$

The dual problem is then to maximize  $g(p)$  with respect to  $p$ . The optimality conditions are the *right Hamilton equations*

$$q(t+1) = D_2 H^+(q(t), p(t+1)) \quad (10a)$$

$$p(t) = D_1 H^+(q(t), p(t+1)). \quad (10b)$$

The system of three equations given by the two Hamilton equations (10) together with the Legendre transform (7) are recognizable as Pontryagin's principle, familiar from optimal control theory. In that setting they are usually formulated in terms of the pseudo-Hamiltonian  $H(q, r, p)$ , where  $q$  is the state,  $r$  is the control and  $p$  is the multiplier<sup>3</sup>.

Apart from minor sign changes, the Legendre transform also occurs as the *Fenchel transform*, which for  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is defined as

$$f^\circ(p) = \sup_x (p^T x - f(x)).$$

The function  $f^\circ$  is also called the *conjugate function* of  $f$ , which is essentially the Hamiltonian. The Fenchel transform plays a central role in Fenchel duality theory [20].

The use of convex duality to derive the Legendre transform shows directly the motivation for the definition of both the Legendre transform and the Hamiltonian. The dual problem is called *Hamilton's principle on the Hamiltonian side*; this is the problem of maximizing  $g(p)$  given by (8). Of course, the approach of convex duality requires the strong assumption of convexity on the action function, although it allows relaxation of smoothness assumptions typically used. However, in order to relax the assumption of convexity, one may of course follow the more usual route via the local theory of Lagrange multipliers [15], which essentially is the spirit of the derivation above. In this case, the conditions for minimality of the action simply become conditions for stationarity of the action. Naturally, the expressions for the

<sup>2</sup> Note that in a differential-geometric formulation the momentum vector  $p$  is a one-form based at  $r$ , so the pseudo-Hamiltonian is a well-defined function on  $Q \times T^*Q$ , whereas it is not so clear where the Hamiltonian  $H(q, p)$  should live.

<sup>3</sup> In optimal control the term *Hamiltonian* is used for the function  $H(q, r, p) = pr - L(r, q)$ , which in mechanics is known as the *pseudo-Hamiltonian*.

Legendre transform are the same, but their interpretation is slightly different. Similarly, one may relax both smoothness and convexity assumptions by following the approach of non-smooth calculus taken by [4].

The important points are that the Hamiltonian arises via duality; the momentum variables are Lagrange multipliers, and the Legendre transform is the map between primal and dual variables that arises from the optimality conditions, i.e., the KKT conditions.

*Left Hamiltonians.* When rendering the augmented action (6) stationary with respect to variations in  $q$ ,  $r$  and  $p$ , we can instead choose to first vary  $q$ , which will give the *left Legendre transform* and *left Hamiltonian*

$$p(t) = -D_1 L(q(t), r(t + 1))$$

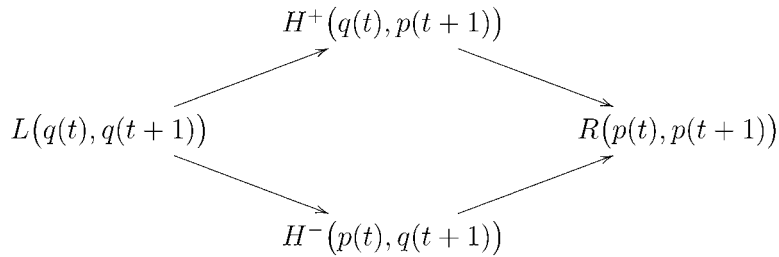
$$H^-(p(t), r(t + 1)) = -\min_q (p(t)^T q(t) + L(q(t), r(t + 1)))$$

and the corresponding *left Hamilton equations*

$$q(t) = -D_1 H^-(p(t), q(t + 1))$$

$$p(t + 1) = -D_2 H^-(p(t), q(t + 1)).$$

Here we have replaced  $r(t)$  by  $q(t)$ ; of course, at optimality they are equal. We can complete the picture by taking another Legendre transform of  $H^-$  or  $H^+$  to obtain a function  $R(p(t), p(t+1))$  with the corresponding evolution equations. The four different functions can be summarized as



with evolution equations

$$\begin{bmatrix} p(t) \\ p(t + 1) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} DL(q(t), q(t + 1))$$

$$\begin{bmatrix} p(t) \\ q(t + 1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} DH^+(q(t), p(t + 1))$$

$$\begin{bmatrix} q(t) \\ p(t + 1) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} DH^-(p(t), q(t + 1))$$

$$\begin{bmatrix} q(t) \\ q(t + 1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} DR(p(t), p(t + 1)).$$

*Continuous time mechanics.* Here we very briefly consider continuous time mechanics to show how the redundant constraints can also be used in this setting to obtain the Hamiltonian side.

Let  $Q$  be a linear finite dimensional configuration space, with tangent bundle  $TQ$ . Given a Lagrangian  $L : TQ \rightarrow \mathbb{R}$ , the *action* of a path  $q(t)$  in  $Q$  is

$$S(q) = \int_0^T L(q(t), \dot{q}(t)) dt. \tag{11}$$

Requiring that the action be stationary with respect to variations in the trajectory yields the familiar *Euler–Lagrange equations*. In the case when the Lagrangian function is convex, one may use this approach to derive the Hamiltonian formulation of mechanics by expressing the variational problem of finding stationary points of the action (11) as

$$\begin{aligned} & \text{minimize} && \int_0^T L(q(t), v(t)) dt \\ & \text{subject to} && \dot{q}(t) = v(t) \quad \text{for all } t \in [0, T]. \end{aligned} \quad (12)$$

Here we would like to minimize over all  $v$  and all  $q$  with the given boundary conditions. The Lagrangian (as the term is used in optimization) is

$$K(q, v, p) = \int_0^T (L(q(t), v(t)) + p(t)^T (\dot{q}(t) - v(t))) dt.$$

The corresponding Lagrange dual function is

$$\begin{aligned} g(p) &= \inf_q \inf_v \int_0^T (L(q(t), v(t)) + p(t)^T (\dot{q}(t) - v(t))) dt \\ &= \inf_q \int_0^T \inf_{v \in \mathbb{R}^{n^*}} (L(q(t), v) + p(t)^T (\dot{q}(t) - v)) dt \\ &= \inf_q S(q, p), \end{aligned}$$

where the dual action  $S(q, p)$  is

$$S(q, p) = \int_0^T (p(t)^T \dot{q}(t) - H(q(t), p(t))) dt. \quad (13)$$

Here we have defined the Hamiltonian by

$$H(q, p) = \inf_{v \in \mathbb{R}^{n^*}} L(q, v) - p^T v;$$

and hence the optimal  $v(t)$  satisfies

$$D_2 L(q(t), v(t)) - p(t) = 0 \quad \text{for all } t \in [0, T].$$

Stationarity with respect to  $q$  and  $p$  gives the *Hamilton equations*

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} D_2 H(q, p) \\ -D_1 H(q, p) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} DH(q, p).$$

The Hamilton equations are exactly equivalent to Pontryagin's minimum principle when the Lagrangian is smooth and strongly convex.

### 3. Sensitivity and its implications

It is a well-known fact that the dual variables  $p$  of the Lagrange dual problem provide a measure of the sensitivity of the primal optimization problem to changes in the constraints. In this section, we show how this sensitivity property leads to reciprocity, symplecticity and momentum maps for discrete mechanics. One may also deduce that standard continuous-time notions of these properties by performing the analogous construction for continuous-time mechanics.

*Sensitivity.* Consider the optimization problem

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && h_i(x) = c_i \end{aligned}$$

with Lagrange multipliers added to give

$$\max_p \min_x (f(x) + p_i^T (h_i(x) - c_i)).$$

If  $x^*(c)$ ,  $p^*(c)$  and  $f^*(c)$  are optimal, then it is easy to show that

$$\frac{\partial f^*(c)}{\partial c_i} = -p_i^*(c),$$

which is the statement of *sensitivity* in optimization.

*Reciprocity.* From the above we have

$$\frac{\partial p_i^*(c)}{\partial c_j} = -\frac{\partial^2 f^*(c)}{\partial c_j \partial c_i}$$

and so we see that the matrix  $\partial p_i^*(c)/\partial c_j$  is thus symmetric, a fact that is termed *reciprocity*.

To apply sensitivity to variational mechanics, consider the value of the action on a solution trajectory,

$$S^*(q_0, q_n) = S(q) \quad \text{if } q \text{ is an optimal solution with } q(0) = q_0, q(n) = q_n.$$

Taking  $p(0)$  and  $p(n)$  to be the Lagrange multipliers enforcing the constraints  $q(0) - q_0 = 0$  and  $q_n - q(n) = 0$ , sensitivity shows that<sup>4</sup>

$$p(0) = -D_1 S^*(q_0, q_n) \quad p(n) = D_2 S^*(q_0, q_n).$$

Then we have the usual form of reciprocity below:

$$\frac{\partial p(i)}{\partial q_j} = -\frac{\partial p(j)}{\partial q_i}.$$

*Symplecticity.* To see how sensitivity and reciprocity relate to symplecticity, consider the two functions:

$$\begin{aligned} g &: (q_0, q_n) \mapsto (-p_0, p_n) \\ f &: (q_0, p_0) \mapsto (q_n, p_n). \end{aligned}$$

The function  $g$  is thus given by  $g = DS^*$ ; so  $Dg = D^2S^*$  is symmetric (which is reciprocity). Furthermore,  $f$  is the flow map that takes initial conditions to final conditions, and we can check<sup>5</sup> that

$$(Df)^T J (Df) = J, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

which is the statement of symplecticity of the flow.

This derivation works for both discrete and continuous time, and for primal (Lagrangian) and dual (Hamiltonian) formulations.

*Generating functions.* Recall that the set of symplectic maps  $f : T^*Q \rightarrow T^*Q$  is equivalent to the set of real valued functions  $F : T^*Q \rightarrow \mathbb{R}$ . This can be seen by considering

$$d[f^*(p_1 dq_1) - p_0 dq_0] = 0,$$

so there exists a  $F(q_0, q_1)$  such that on the graph of  $f$

$$p_1 dq_1 - p_0 dq_0 = dF(q_0, q_1) = D_1 F(q_0, q_1) dq_0 + D_2 F(q_0, q_1) dq_1$$

<sup>4</sup> In mechanics it is standard to derive these relations directly from the Euler–Lagrange equations (see, e.g., [16] or [18]).

<sup>5</sup> To prove that  $Dg = (Dg)^T$  implies  $(Df)^T J (Df) = J$  requires either a straightforward, if tedious, calculation in coordinates, or one can observe that  $0 = d^2 S^* = d(-p_0^i dq_0^i + p_1^i dq_1^i) = dq_0^i \wedge dp_0^i - dq_1^i \wedge dp_1^i$  which is the equivalent statement in differential-geometric notation.

and so  $f$  is defined implicitly from  $F$  by

$$p_0 = -D_1 F(q_0, q_1) \quad p_1 = D_2 F(q_0, q_1).$$

Here  $F$  is called a *generating function of the first kind* [9] and we can see that it is exactly a discrete Lagrangian, with  $f$  being the corresponding evolution, as shown in [18]. There are four types of classically recognized generating functions, based on the four coordinate choices  $(q_0, q_1)$ ,  $(q_0, p_1)$ ,  $(p_0, q_1)$ , and  $(p_0, p_1)$  for the graph of  $f$ .

These are exactly the same as the four functions  $L$ ,  $H^+$ ,  $H^-$  and  $R$ , and the generating function equations are exactly the corresponding Hamilton equations. This implies that all symplectic integrators can be derived from a discrete time Lagrangian or Hamiltonian. Generating functions were amongst the first techniques used to construct symplectic integrators (see, for example, [21], [3], [6] and [7]), and were also used in the analysis of symplectic schemes, such as in [8].

*Symmetries and momentum maps.* Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$  acting on  $Q$ . We say that  $G$  is a *symmetry* of a Lagrangian if

$$L(q(t), q(t+1)) = L(g \cdot q(t), g \cdot q(t+1))$$

for all  $g \in G$ , or the equivalent infinitesimal statement

$$D_1 L(q(t), q(t+1)) \cdot \xi(q(t)) + D_2 L(q(t), q(t+1)) \cdot \xi(q(t+1)) = 0$$

for all  $\xi \in \mathfrak{g}$ . This implies that the action  $S(q)$  is symmetric, and so too is  $S^*(q_0, q_n)$ . We thus have

$$D_1 S^*(q_0, q_n) \cdot \xi(q_0) + D_2 S^*(q_0, q_n) \cdot \xi(q_n) = 0$$

and using sensitivity gives *Noether's theorem* of momentum conservation

$$p_0 \cdot \xi(q_0) = p_n \cdot \xi(q_n).$$

Variational integrators thus preserve momentum maps.

#### 4. Separable optimization problems and Hamilton–Jacobi theory

The action sum (2) has the special property that it is the sum of functions whose dependence on the optimization variables  $x(0), \dots, x(n)$  is local. This causes the optimization to be *separable*. An important consequence of this is the associated Hamilton–Jacobi theory, where the action sum is split into separate terms, and the optimal action is expressed in terms of the optimal values of each of the terms.

To see this, consider the action evaluated on a trajectory as a function of two times and two positions

$$S^*(t, q_t, \ell, q_\ell) = S(q),$$

where  $q$  is a solution satisfying  $q(t) = q_t$  and  $q(\ell) = q_\ell$ . For any  $m$  with  $t < m < \ell$  we have

$$S^*(t, q_t, \ell, q_\ell) = S^*(t, q_t, m, q_m^*) + S^*(m, q_m^*, \ell, q_\ell)$$

where  $q_m^*$  renders the right-hand side stationary. In the particular case that  $m = t + 1$  and  $\ell = n$  this is the discrete time *Hamilton–Jacobi equation*:

$$S^*(t, q_t, n, q_n) = L(q_t, q_{t+1}^*) + S^*(t + 1, q_{t+1}^*, n, q_n),$$

where  $q_{t+1}^*$  renders the right-hand side stationary. By taking the limit  $\Delta t \rightarrow 0$  this recovers the continuous time Hamilton–Jacobi equation:

$$-D_1 S^*(t, q_t, T, q_T) = H(q, -D_2 S^*(t, q_t, T, q_T)).$$



This equation can also be derived by realizing that the flow map  $f^{T-t} : (q_t, p_t) \mapsto (q_T, p_T)$  is symplectic, and hence has a generating function  $S^*(t, q_t, T, q_T)$ , and then computing the equation that describes the time evolution of  $S^*$ .

One may also construct a Hamilton–Jacobi theory for multisymplectic problems, again, due to the separable structure of the action sum. The Hamiltonian is important to the Hamilton–Jacobi theory, and one can recognize it as playing the same role as that of the conjugate function in separable optimization problems.

### 5. Symplectic integrators

While classical generating function theory tells us that every symplectic integrator arises from a discrete Lagrangian, it turns out that many well-known methods arise very naturally from a variational formulation. These include Verlet, conservative Newmark and Runge–Kutta methods (see [18] for details). The symplectic Runge–Kutta methods (identified by [12] and [23]) have a particularly neat variational derivation, as shown by [22].

To derive numerical integrators using discrete Hamiltonians, it is important to recognize that it is necessary for the discrete action sum (2) to approximate the continuous action (11) [18, 11]. On the Hamiltonian side this means that (9) should approximate (13). In particular, a right Hamiltonian should satisfy

$$H^+(q_0, p_1) \approx p_1^T q_1 - \int_0^{\Delta t} (p(t)^T \dot{q}(t) - H(q(t), p(t))) dt$$

As a simple example of this, consider the first-order approximation

$$\begin{aligned} H^+(q_0, p_1) &= p_1^T (q_0 + \Delta t D_2 H(q_0, p_1)) - \Delta t (p_1^T D_2 H(q_0, p_1) - H(q_0, p_1)) \\ &= p_1^T q_0 + \Delta t H(q_0, p_1). \end{aligned}$$

Computing the Hamilton equations (10) gives

$$\begin{aligned} q_1 &= D_2 H^+(q_0, p_1) = q_0 + \Delta t D_2 H(q_0, p_1) \\ p_0 &= D_1 H^+(q_0, p_1) = p_1 + \Delta t D_1 H(q_0, p_1). \end{aligned}$$

In the special case  $H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q)$ , the Hamilton equations become

$$q_1 = q_0 + \Delta t M^{-1} P_1 \quad p_1 = p_0 - \Delta t DV(q)$$

and we recognize the first-order symplectic-Euler method [11]. Higher order methods result from taking higher order approximations to the action.

Note that while the integrator resulting from the discrete Hamiltonian is symplectic, it certainly will not preserve either the continuous-time Hamiltonian  $H(q, p)$  or the discrete-time Hamiltonian  $H^+(q_0, p_1)$  exactly, although both will be approximately conserved as dictated by backward error analysis [11].

### 6. Conclusions

As we have seen above, augmenting with additional constraints and using Lagrange duality theory provides a canonical choice of a Hamiltonian counterpart to discrete Lagrangian mechanics. The dual nature of these formulations is familiar from discrete optimal control theory and generating function theory, and provides a natural choice for constructing discrete Hamiltonians and discrete Hamilton equations. In addition, the sensitivity analysis of optimization theory provides a direct method for establishing symplecticity of the discrete Hamilton system and for determining its momentum maps.

The theory presented above extends immediately to a discrete multisymplectic setting, where we now have a mesh of discrete variables rather than a line in time. The duality theory extends naturally, but there are now more than two discrete Hamiltonians, as each element in the mesh has one discrete Hamiltonian per node.

A simple application of multisymplectic theory is for solving elliptic systems, for example, when solving for the equilibrium configuration of an elastic system. This problem has a natural variational formulation, and one may perform the dual construction as in the hyperbolic case. If one is solving this problem over a fixed domain with appropriate boundary conditions, then the dual structure gives rise to a natural class of gradient algorithms via the *dual-decomposition*. Both primal and dual variants of these algorithms are decentralized, and proceed by alternately solving the primal and dual system to compute the gradient. On the primal side, this corresponds to associating configuration variables with elements. In one dimension, one associates momentum variables with nodes in between adjacent elements, and the configuration of two adjacent elements is used to compute left and right momenta, which are passed to the nodes. The nodes in turn update their configuration using these momenta. The dual algorithm is similar; in this case nodes have a single associated momentum variable and elements each have left and right configuration variables. Both of these algorithms result from direct application of the gradient method, either to Hamilton's principle or to the dual variational problem, and the decentralized structure arises immediately from the separable structure of the action. It is interesting to note also that a very similar decentralized gradient algorithm is used in congestion control of networks [14].

Such algorithms become interesting in multisymplectic models, where one has an array of elements over independent variables representing space-time. If the associated discretization is an irregular mesh, then each element may be responsible for a different number of momentum and configuration variables, depending on the local structure of the mesh. The resulting algorithms are related to the asynchronous variational integrators described in [13]. Understanding the dual structure of the discrete mechanics here may also allow application to distributed optimal control systems with large arrays of actuators and sensors.

## References

- [1] Bertsekas D P 2005 *Dynamic Programming and Optimal Control* vol 1, 3rd edn (Belmont, MA: Athena Scientific)
- [2] Boyd S and Vandenberghe L 2004 *Convex Optimization* (Cambridge: Cambridge University Press)
- [3] Channell P J and Scovel C 1990 Symplectic integration of Hamiltonian systems *Nonlinearity* **3** 231–59
- [4] Clarke F H 1990 *Optimization and Nonsmooth Analysis* (Philadelphia: SIAM)
- [5] de León M, Martín de Diego D and Santamaría-Merino A 2003 Discrete variational integrators and optimal control theory *Preprint math/0311169*
- [6] Feng K 1986 Difference-schemes for hamiltonian-formalism and symplectic-geometry *J. Comput. Math.* **4** 279–89
- [7] Feng K, Wu H M, Qin M Z and Wang D L 1989 Construction of canonical difference-schemes for Hamiltonian-formalism via generating-functions *J. Comput. Math.* **7** 71–96
- [8] Ge Z and Marsden J M 1988 Lie–Poisson integrators and Lie–Poisson Hamilton–Jacobi theory *Phys. Lett. A* **133** 134–9
- [9] Goldstein H 1980 *Classical Mechanics* 2nd edn (Reading, MA: Addison-Wesley)
- [10] Gonzalez O 1996 Time integration and discrete Hamiltonian systems *J. Nonlinear Sci.* **6** 449–67
- [11] Hairer E, Lubich C and Wanner G 2002 *Geometric Numerical Integration* (Berlin: Springer)
- [12] Lasagni F M 1988 Canonical Runge–Kutta methods *Z. Angewandte Math. Phys.* **39** 952–3
- [13] Lew A, Marsden J E, Ortiz M and West M 2003 Asynchronous variational integrators *Arch. Ration. Mech. Anal.* **167** 85–146
- [14] Low S and Lapsley D E 1999 Optimization flow control. I. Basic algorithm and convergence *IEEE/ACM Trans. Netw.* **7** 861–74

- [15] Luenberger D G 1969 *Optimization by Vector Space Methods* (New York: Wiley)
- [16] Marsden J E and Ratiu T 1999 *Introduction to Mechanics and Symmetry* vol 17, 2nd edn (Berlin: Springer)
- [17] Rowley C W and Marsden J E 2002 Variational integrators for degenerate Lagrangians, with application to point vortices *Proc. 41st IEEE Conf. Decision and Control* 1521–7
- [18] Marsden J E and West M 2001 Discrete mechanics and variational integrators *Acta Numerica* vol 10 (Cambridge: Cambridge University Press) DOI <http://dx.doi.org/10.1017/S096249290100006X>
- [19] McLachlan R I, Quispel G R W and Robidoux N 1999 Geometric integration using discrete gradients *Phil. Trans. R. Soc. A* **357** 1021–45
- [20] Rockafellar R T 1974 *Conjugate Duality and Optimization* (Philadelphia: SIAM)
- [21] Ruth R D 1983 A canonical integration technique *IEEE Trans. Nucl. Sci.* **30** 2669–71
- [22] Suris Y 1990 Hamiltonian methods of Runge-Kutta type and their variational interpretation *Math. Simul.* **2** 78–87
- [23] Suris Y B 1989 The canonicity of mappings generated by Runge-Kutta type methods when integrating the system  $\ddot{x} = -\partial u/\partial x$  *USSR Comput. Math. Math. Phys.* **29** 138–44